



Climate Change

EQC Status Overview

Francisco Doblas-Reyes^{1,2}

1 Barcelona Supercomputing Center (BSC)

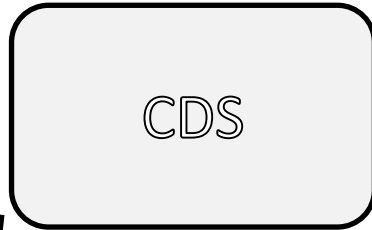
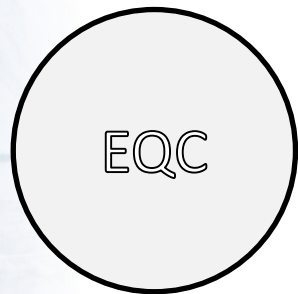
2 Institució Catalana de Recerca i Estudis
Avançats (ICREA)





Climate
Change

EQC: Engaged and future activities



Quality assurance for seasonal forecasts

Quality assurance framework for earth observations

Quality assurance for climate projections

Quality assessment of ECV products and reanalyses

Sectoral gap analysis and user requirements

EQC of operational SIS

- Ensures C3S is state-of-the-art
- Identifies gaps in the Service
- Bridges Copernicus with Research Agenda in Europe (e.g. H2020, national research projects)
- Monitors continually, quality of C3S products and services
- “Quality Assurance” body
- Contributes and develops URDB/SES/etc documents

J.N. Thépaut (ECMWF)



Climate
Change

Challenges of the EQCs

- **Address adaptation:** it must provide information for all kind of services including adaptation, taking advantage that many users are already familiar with the climate-change problem.
- **Provide consistency:** it must build trust, ensuring a high degree of coherence across products, underlying data sets, processing methods, communication, etc.
- **Provide innovation:** it should make recent developments from research operational to answer real-world issues.
- **Address efficiency:** the EQC information should be timely, e.g. respond to users' queries with a delay as short as possible, which imposes conditions on the algorithms considered.
- **Define the target:** data, products, communication, etc.



Climate
Change

Framework for collaboration

C3S EQC Workshop
12-14 June 2017



SAVE THE DATE

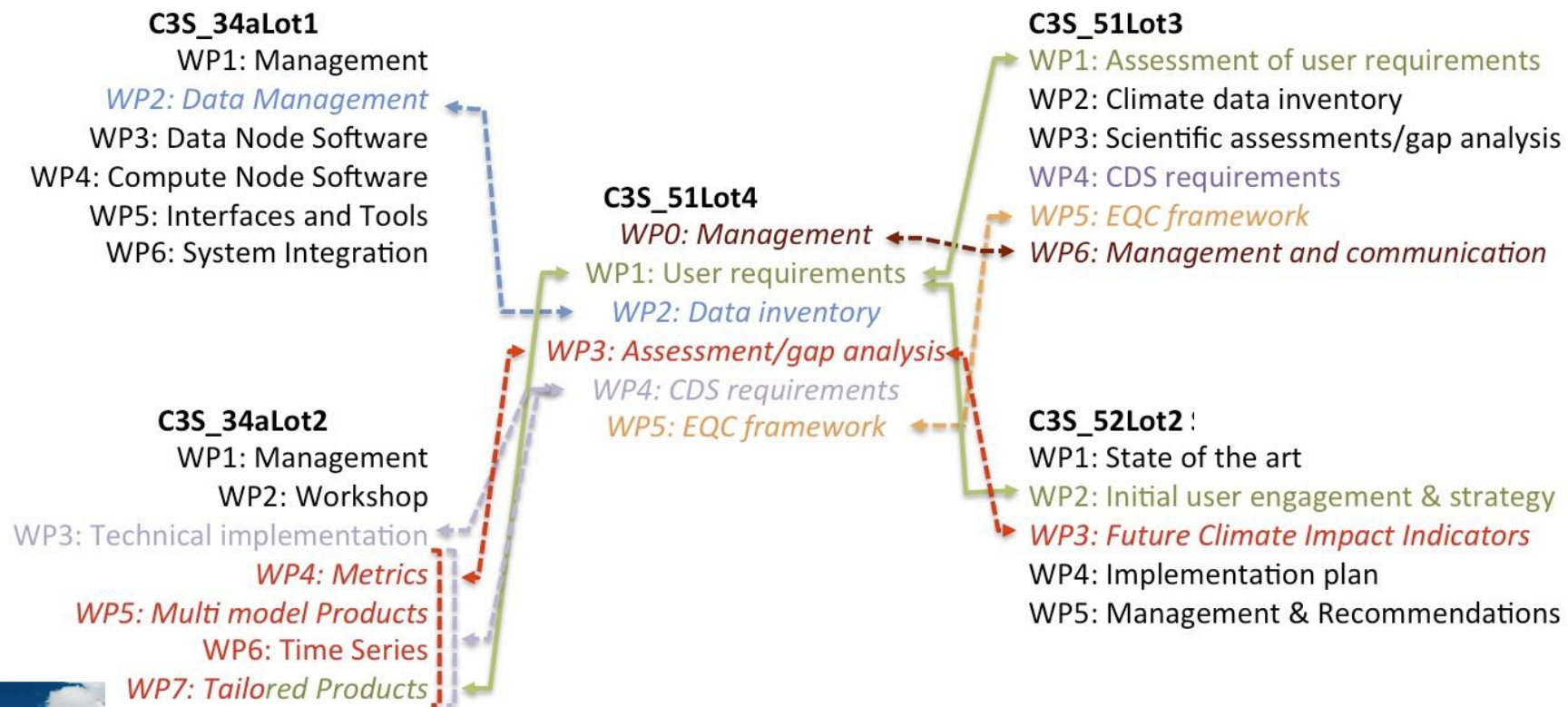
Copernicus Climate Change Service –
Evaluation and Quality Control workshop

Date: 12-14 June 2017
Location: Barcelona Supercomputing Center, Barcelona, Spain
Register:
<https://climate.copernicus.eu/events/evaluation-and-quality-control-workshop>

C3S EQC workshop

The event will focus on:

- Progress to date for each of the contracts (Observations, Seasonal, Climate and User Requirements)
- Identification of any gaps that we should address
- How the EQC function should work once C3S is operational



H. Gregow (FMI)

Some examples:

- Observations are expected to provide accuracy estimates.
- Reanalyses should come along with a validation against independent observations.
- Forecasts have no real value without an estimate of quality based on past performance.
- Projections require a scientific validation of the models used (e.g. IPCC WGI Ch. 9).

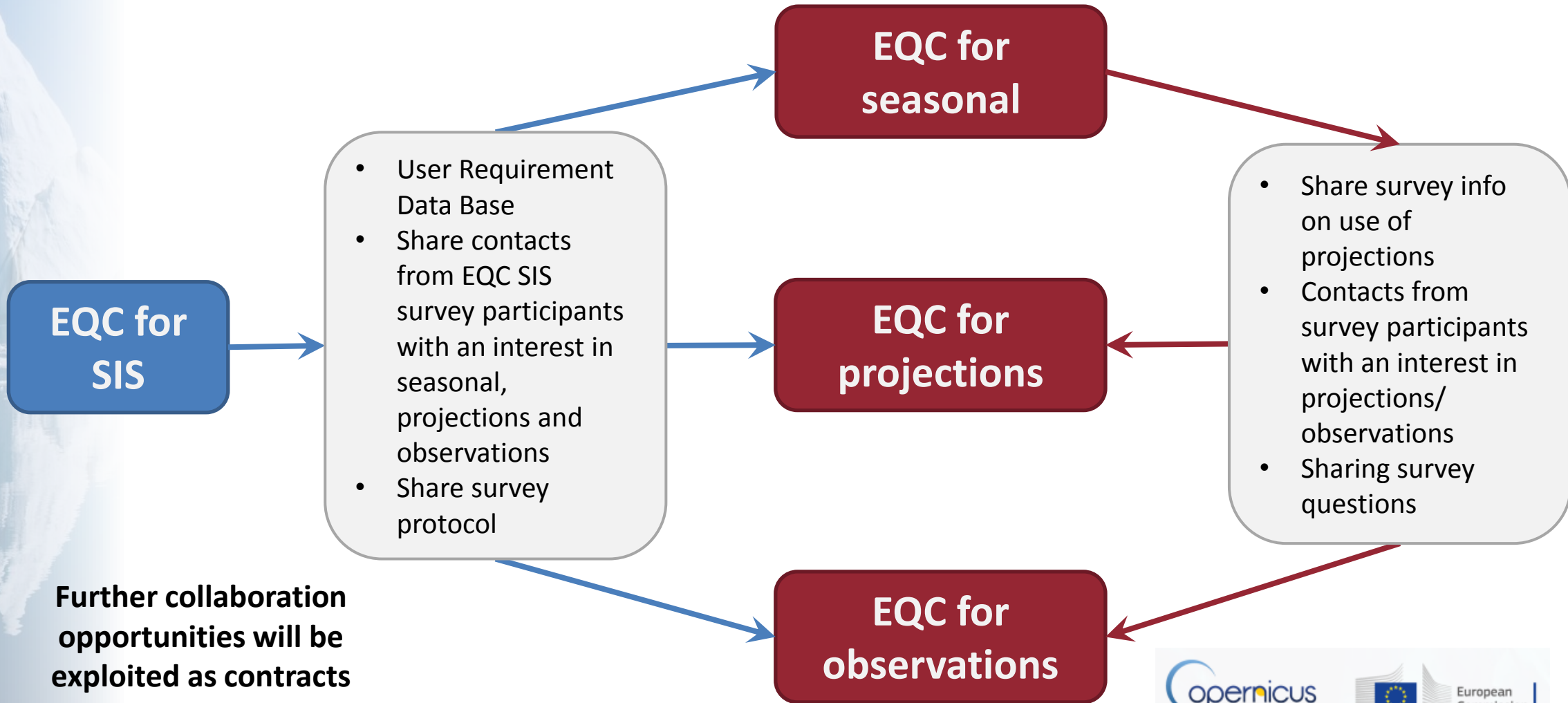
But quality information is much more complex than this.



Climate
Change

Collaboration between EQCs

User engagement and assessment of needs



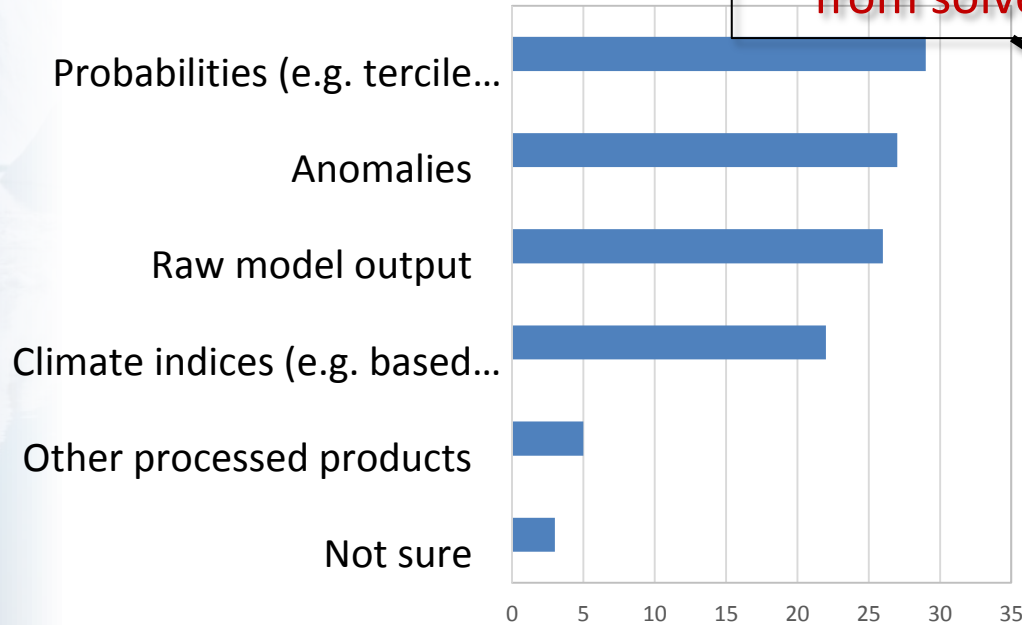
Further collaboration opportunities will be exploited as contracts develop further

M. Soares (Univ. Leeds)

Users' requests

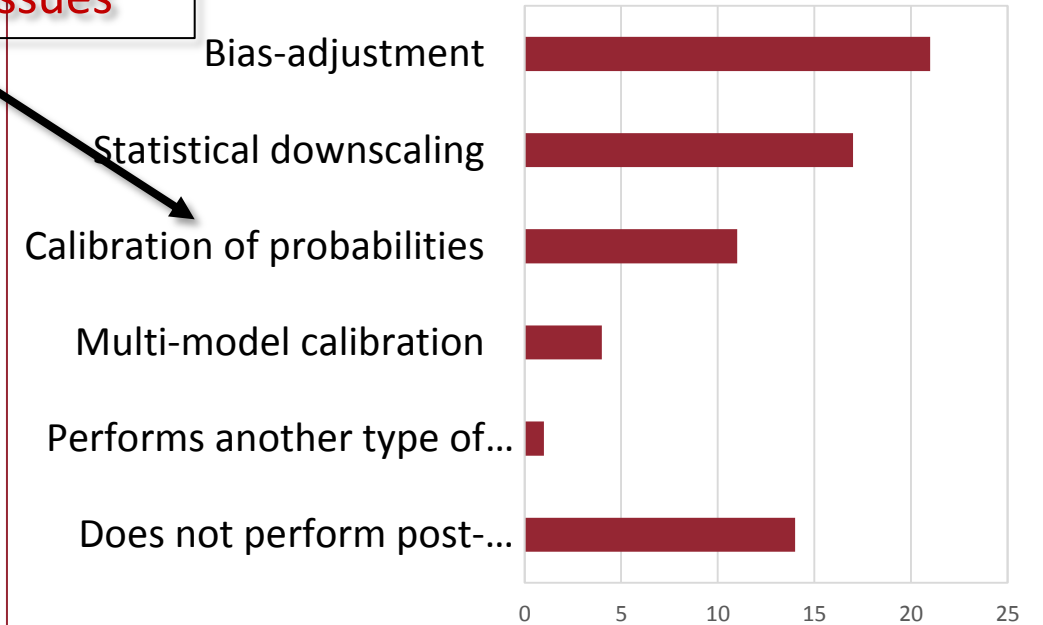
Results from a [survey](#) where 42 out of 53 respondents receive seasonal forecast information, with a large majority of NMHSs.

"What kind of data from global seasonal forecast models do you use?"



Note that these are far from solved issues

"What type of adjustment post-processing do you perform on the SCF data before using it?"





Climate
Change

Data inventories

Evaluation matrices for the different aspects

		GCM-RCM chains																					
		EUR-44	RCM-GCM chains																	Total number			
			CNRM-CM5	CNRM-CM5	MPI-ESM-LR	EC-EARTH	EC-EARTH	HadGEM2-ES	CanESM2	CNRM-CM5	CSIRO-Mk3-6-0	EC-EARTH	IPSL-CM5A-MR	MIROC5	HadGEM2-ES	MPI-ESM-LR	NorESM1-M	GFDL-ESM2M	MPI-ESM-LR		IPSL-CM5A-MR		
			ALADIN52	ALADIN53	CCLM4-8-17	HIRHAM5	RACMO22E	RCA4											REMO2009		WRF331F		
Type of grid	Output frequency																						
EUR-44	3-hourly																						0
EUR-44i	6-hourly																						0
EUR-11	daily		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	18
EUR-11i	monthly		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	18
	seasonal		x		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x		16
	invariant (fx)		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	18
		EUR-11	RCM-GCM chains																	Total number			
			CNRM-CM5	CNRM-CM5	EC-EARTH	HadGEM2-ES	MPI-ESM-LR	EC-EARTH	EC-EARTH	HadGEM2-ES	CNRM-CM5	EC-EARTH	IPSL-CM5A-MR	HadGEM2-ES	MPI-ESM-LR	MPI-ESM-LR	IPSL-CM5A-MR						
			ALADIN53	CCLM4-8-17			HIRHAM5	RACMO22E		RCA4								REMO2009	WRF331F				
Spatial resolution	Output frequency																						
0.44°	3-hourly										x	x	x	x	x							5	
0.11°	6-hourly										x	x	x	x	x							5	
	daily		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					15	
Land area fraction available	monthly		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					15	
Orography available	seasonal			x	x	x	x	x	x	x	x	x	x	x	x	x						13	
Pressure level data available	invariant (fx)		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					15	
Temporal resolution available on both EUR-44 and EUR-11																							

Temporal resolution available on both EUR-44 and EUR-11

Multiplied criteria:
resolution, data frequency,
available variables and scenarios

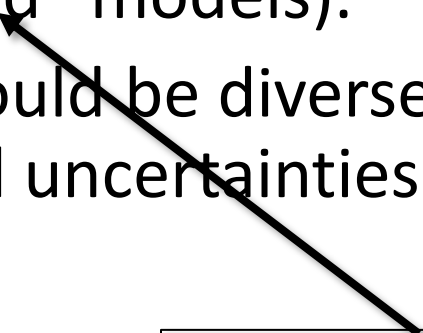
Narrow ensemble

Proposal to CDS

Once the data are identified and available, a critical question is which of the available data will be used as the data use will depend on, among other things, the quality information.

C3S51_Lot4 considers two criteria:

- Inadequate sources for the user's application should be omitted (e.g. discard "bad" models).
- The sources retained should be diverse enough to still adequately represent all uncertainties (e.g. has the solution enough spread?)



What does "bad"
mean?

Definitions are important:

- **Diagnostics** can be derived from any geophysical data set, independently of any reference (e.g. Arctic sea-ice area).
- **Metrics** reflect the agreement of a diagnostic from a system with respect to the same diagnostic computed from a reference.
 - Performance: RMSE of Arctic sea-ice area in a historical run.
 - Predictability: spread of Arctic sea-ice area in a multi-model.
 - Forecast quality: explained variance of Arctic sea-ice area.
 - Process-based: heat budget analysis of Arctic sea ice.
- A **constraint** is a metric applied to an ensemble of simulations displaying relationships between two diagnostics, one of which can be observed.
- **Diagnosis** is an integrated statement for an EQC purpose. It involves the use of diagnostics and different metrics, together with prior knowledge about the system itself and its underlying physics.
- The **CRISTO** framework for evaluation might help: **completeness, rationale, interpretability, stability, transparency, observability.**

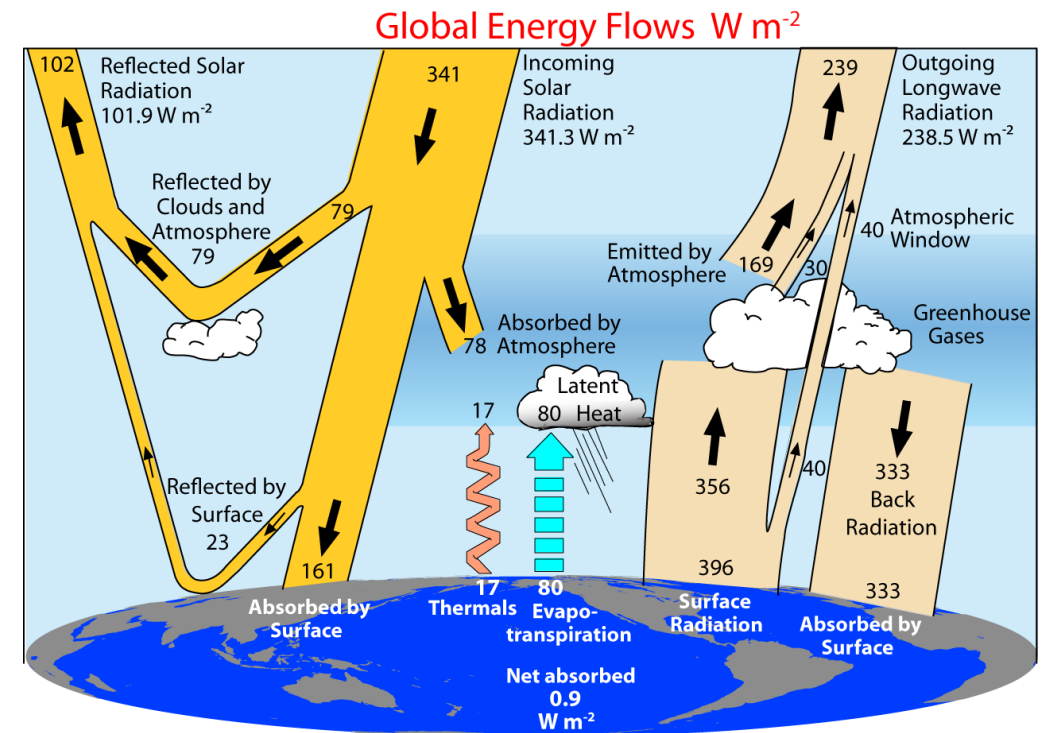
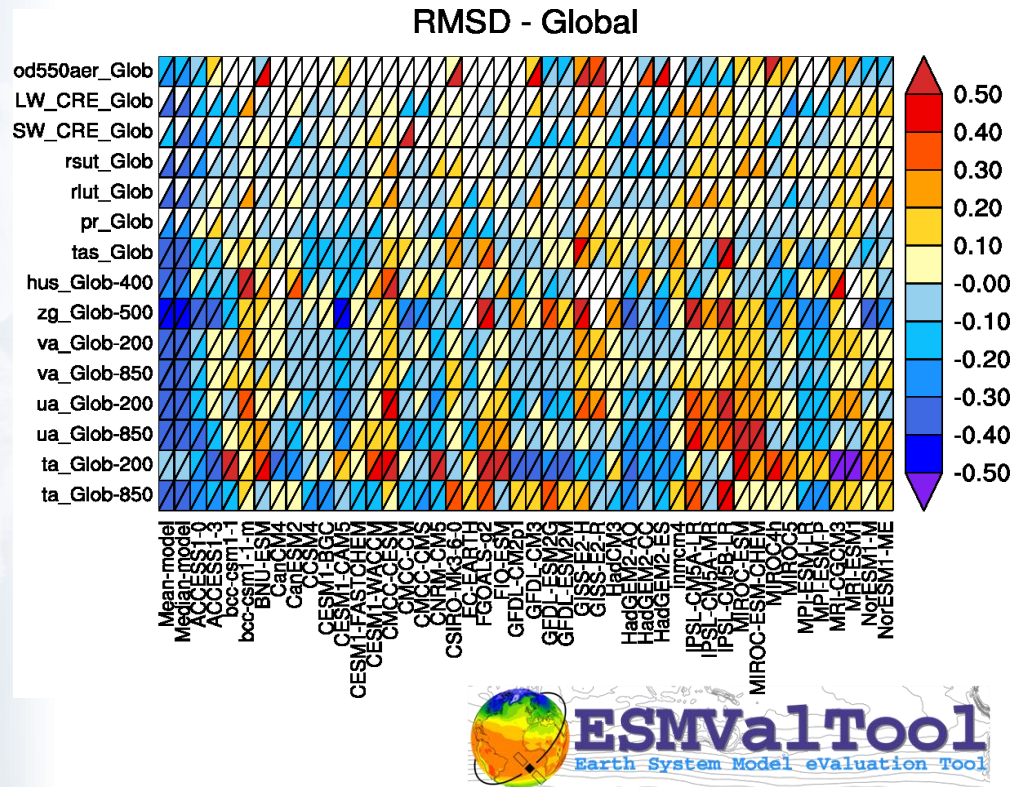




Climate
Change

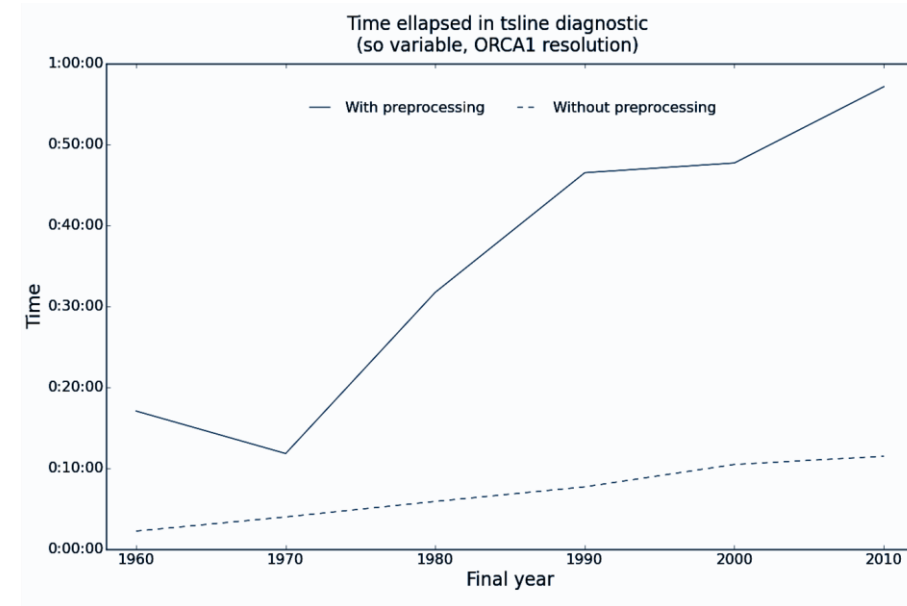
Scientific assessment and gap analysis

Diagnostics and metrics for climate projections: ESMValTool, CVDP (NCAR), AutoAssess (Met Office), s2dv (BSC), Freva (FUB), PMP (PCMDI).



The example of ESMValTool:

- Recoding all the diagnostics and metrics that it contains is prohibitive and prone to errors. **Benefit from the community.**
- A rewrite of the backend is taking place as we speak: migration to python, communication in-memory, common pre-processing, parallelisation, etc.
- Data provenance to be moved from being stored in the output file plus a logfile to full provenance approach.
- But waiting for C3S to define requirements !!

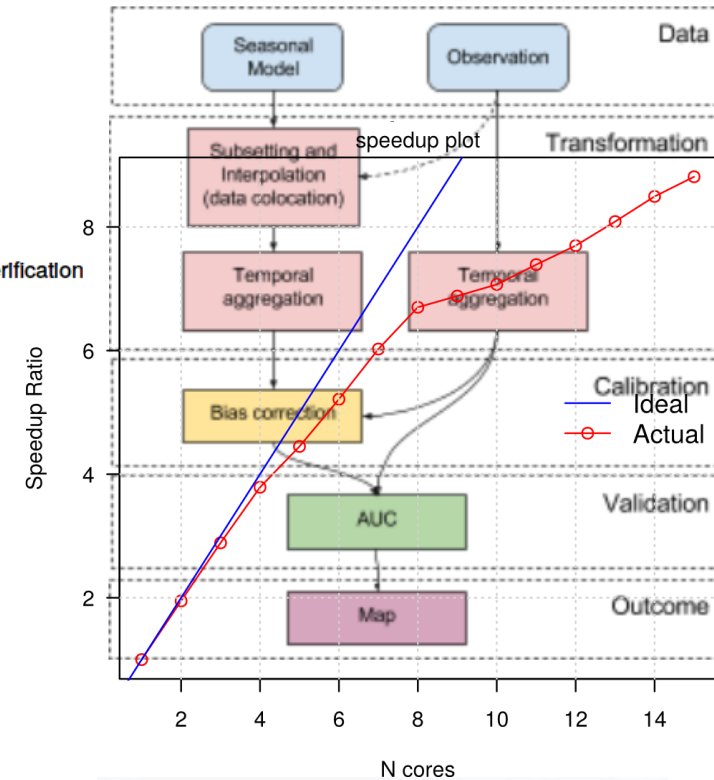
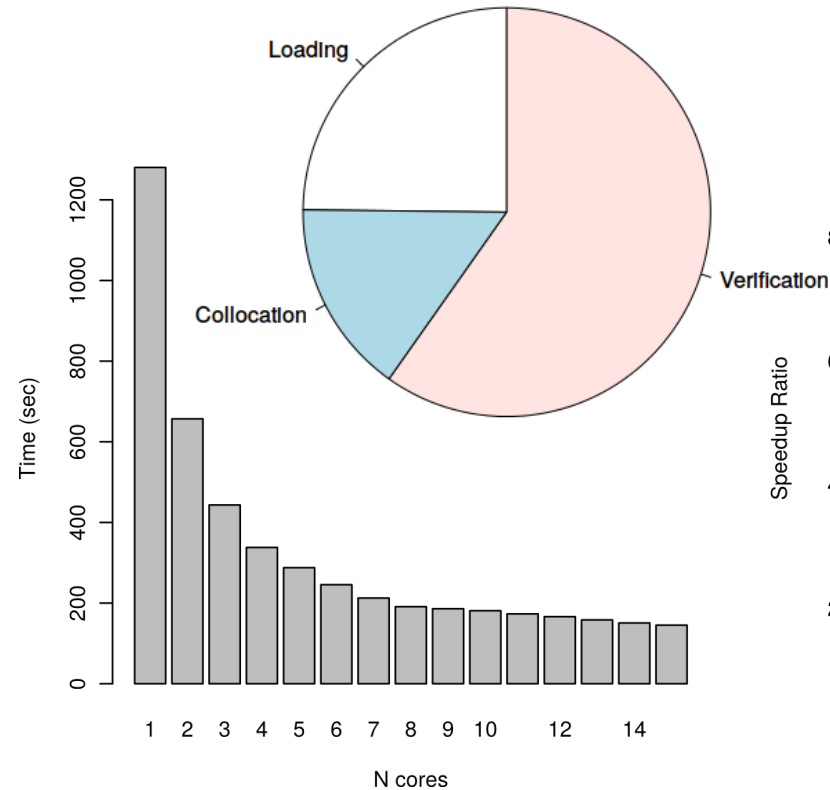
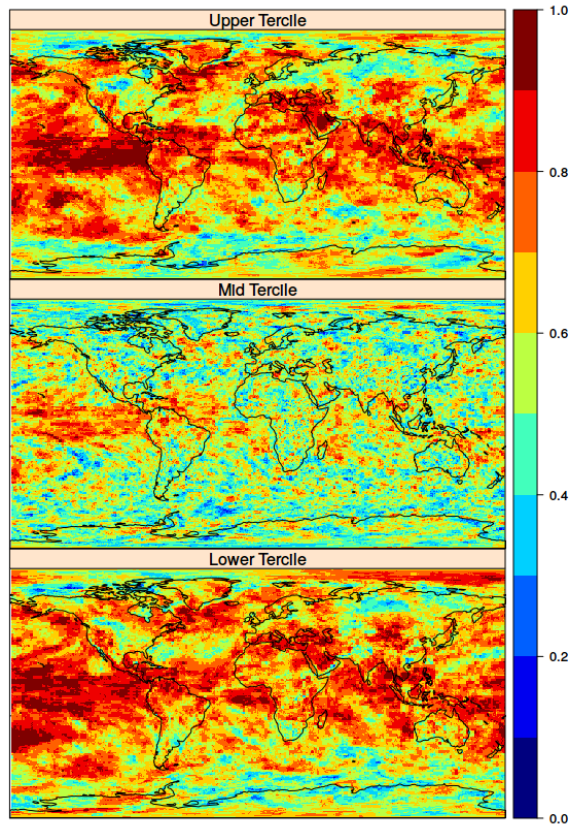




Climate
Change

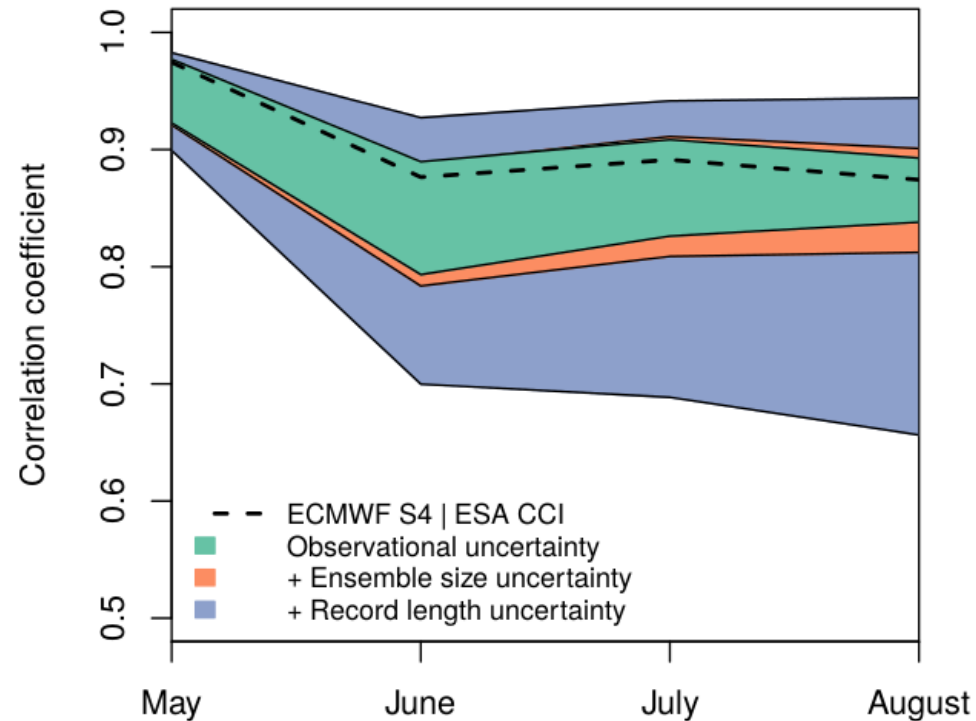
Scientific assessment and gap analysis

More on performance: example of multi-model seasonal verification, with performance analysis of a ROC area estimate using loadeR, SpecsVerification and easyVerification



Making the most of C3S: bringing in observational and reference uncertainty in the EQC process.

Niño 3.4 SST correlation of the ensemble mean for EC-Earth3.1 (T511/ORCA025) predictions with ERAInt and GLORYS2v1 ics, and BSC sea-ice reconstruction started every May over 1993-2009



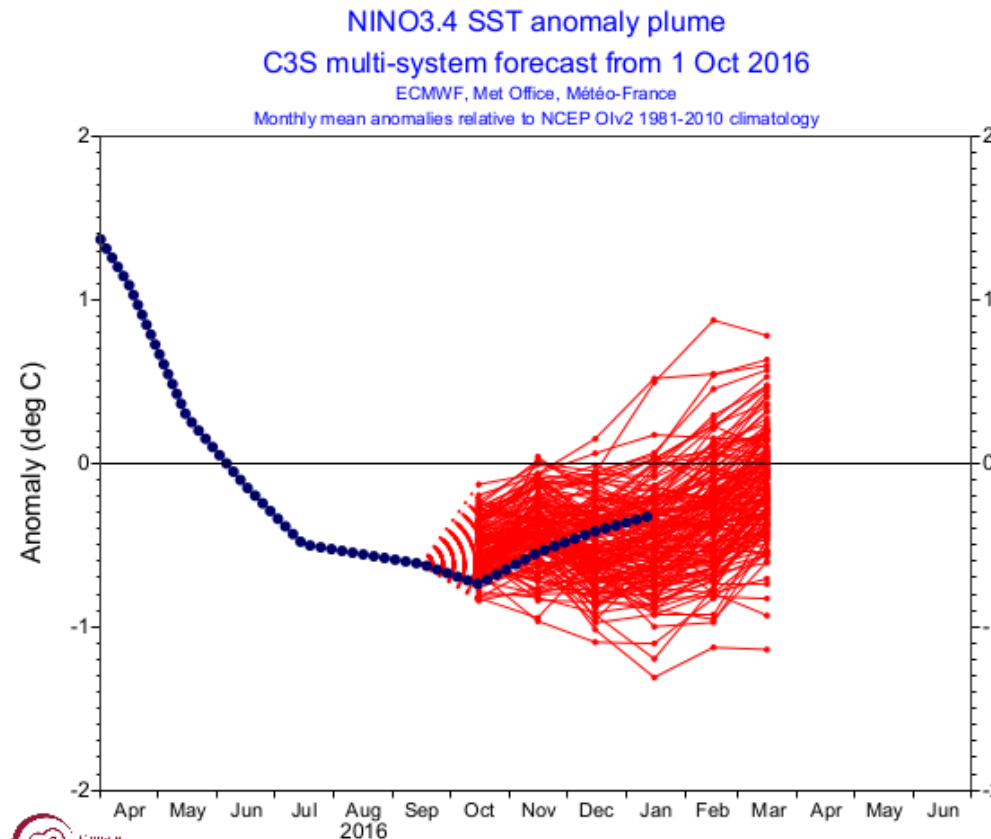


Climate
Change

CDS requirements and EQC framework

How to identify data/products to ensure a minimum quality?

- Reproducibility: ability of an entire process to be duplicated.
- Traceability: ability to verify the history, location, or application of an item by means of documented recorded identification.



How to identify data/products to ensure a minimum quality?

- Reproducibility: ability of an entire process to be duplicated.
- Traceability: ability to verify the history, location, or application of an item by means of documented recorded identification.

Seasonal forecasts of wind speed

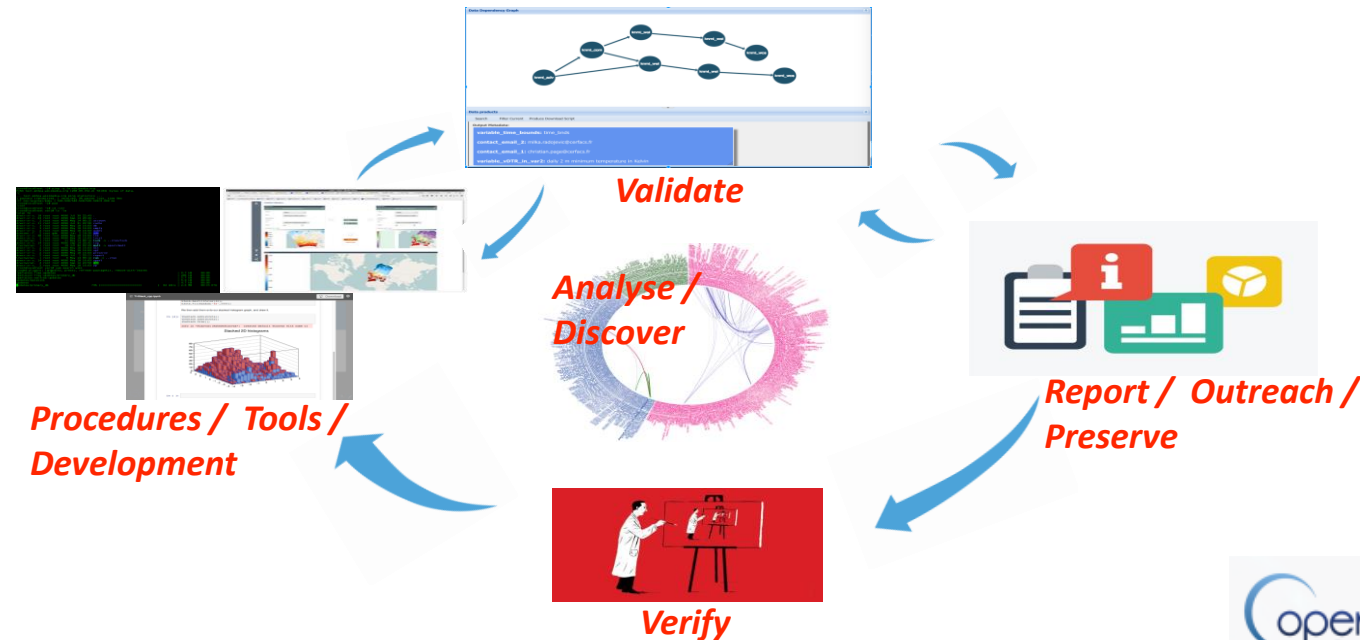


Climate
Change

CDS requirements and EQC framework

How to identify data/products to ensure a minimum quality?

- Generalised metadata and provenance information are key elements of all the components of the service.
- Two approaches for product provenance are under discussion: S-PROV and Resource Description Framework (RDF). They are not mutually exclusive.



A. Spinuso (KNMI)

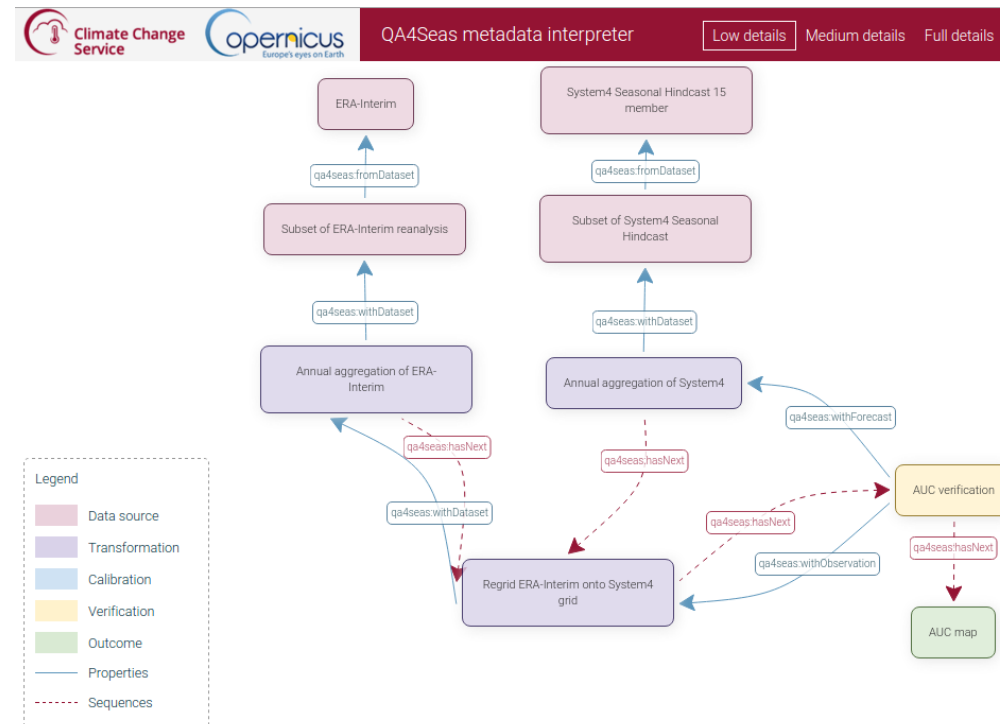


Climate
Change

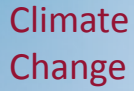
CDS requirements and EQC framework

How to identify data/products to ensure a minimum quality?

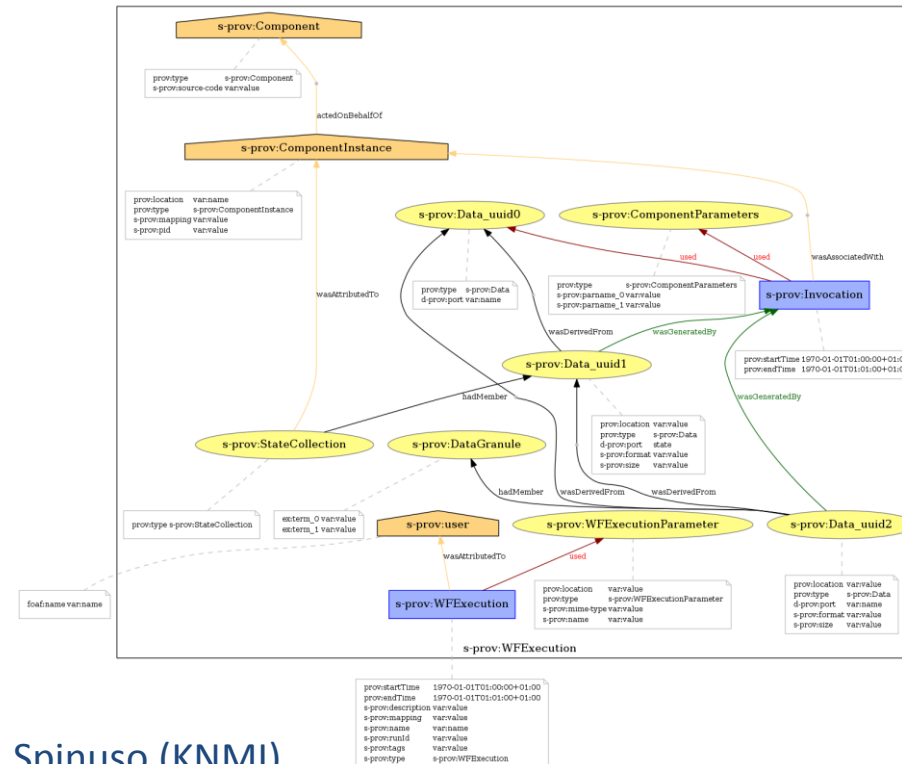
- The [RDF-based approach](#) aims at the reproducibility of objects (NetCDF file, image) with human and machine-readable solution.
- It uses a semantic metadata model that builds the vocabularies on existing initiatives (e.g. VALUE for downscaling).



D. San Martín (PREDICTIA)



- S-PROV represents the relationships occurring between the players of a data-intensive computation in a scientific domain.
- Uses W3C PROV as data model and S-PROVFlow for reproducibility as a service.



Provenance and metadata challenges:

- Engage the (expert) users.
- Define the level of granularity to describe the objects.
- Inform about and display different levels of abstraction.
- Define the curation of elements other than raw data.
- **Which components of the C3S are involved and where does the governance reside?**

S u m m a r y

- EQC is user driven, but not all users are feeding in yet.
- Data inventories help identifying gaps.
- EQC information is not neutral, precise definitions are necessary.
- Existing packages are an invaluable source of solutions, but should be considered within a framework.
- Handling metadata and provenance information require a generic, common approach for all the EQC (and other components) work in C3S.